

2. Covariance

Covariance

Claim Term	CMU's Construction	Marvell's Construction
<p>covariance</p> <p>'839 Patent Claims 11, 16, 19, 23 '180 Patent Claim 6</p>	<p>none; see "noise covariance matrices"</p> <p>CMU Brf. at 27</p>	<p>the expected (mean) value of the product of $(r_i - m_i)$ and $(r_j - m_j)$, where r_i and r_j are observed signal samples (at time i and time j, respectively) and m_i and m_j are the expected (mean) values of the samples (at time i and time j, respectively) (i.e., $E[(r_i - m_i)(r_j - m_j)]$).</p> <p>Marvell Brf. at 21-24</p>

- Dispute
 - ▶ Should "covariance" have its ordinary meaning (Marvell) or be read out of the claim (CMU)?

Claim Language

- Covariance is used to calculate the correlation-sensitive branch metrics

US 6,201,839 B1

15
outputting said branch weight.

11. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols stored on a high density magnetic recording device, comprising the steps of:

(a) performing a Viterbi-like sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;

(b) outputting a delayed decision on the recorded symbol;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and said delayed decisions;

(e) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and

(f) repeating steps (a)-(e) for every new signal sample.

12. The method of claim 11 wherein said Viterbi-like sequence detection is performed using a PRML algorithm.

13. The method of claim 11 wherein said Viterbi-like sequence detection is performed using an FDTSD algorithm.

14. The method of claim 11 wherein said Viterbi-like sequence detection is performed using an RAM-RSE algorithm.

15. The method of claim 11 wherein said Viterbi-like sequence detection is performed using an MDFE algorithm.

16. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising the steps of:

(a) performing a Viterbi-like sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;

(b) outputting a delayed decision on the transmitted symbol;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and said delayed decisions;

(e) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and

(f) repeating steps (a)-(e) for every new signal sample.

17. The method of claim 16 wherein said channel is nonstationary noise.

18. The method of claim 16 wherein said channel is nonstationary signal dependent noise.

19. A detector circuit for detecting a plurality of signal samples read from a recording medium, comprising:

a Viterbi-like detector circuit, said Viterbi-like detector circuit for producing a plurality of delayed decisions and a plurality of delayed signal samples; and

a noise statistics tracker circuit responsive to said Viterbi-like detector circuit for updating a plurality of noise covariance matrices in response to said delayed decisions and said delayed signal samples; and

a correlation-sensitive metric computation update circuit responsive to said noise statistics tracker circuit for recalculating a plurality of correlation-sensitive branch metrics from said noise covariance matrices, said branch metrics output to said Viterbi-like detector circuit.

19. A detector circuit for detecting a plurality of data from a plurality of signal samples read from a recording medium, comprising:

a Viterbi-like detector circuit, said Viterbi-like detector circuit for producing a plurality of delayed decisions and a plurality of delayed signal samples from a plurality of signal samples;

a noise statistics tracker circuit responsive to said Viterbi-like detector circuit for updating a plurality of noise covariance matrices in response to said delayed decisions and said delayed signal samples; and

a correlation-sensitive metric computation update circuit responsive to said noise statistics tracker circuit for recalculating a plurality of correlation-sensitive branch metrics from said noise covariance matrices, said branch metrics output to said Viterbi-like detector circuit.

See '839 Patent Claims 11, 16, 19, 23; '180 Patent Claim 6

Claim Language

- “Covariance” is used independently without “noise covariance matrices”

13 shows the performance of the PR4 detectors at this density. FIG. 9 is similar to FIG. 7, except that the error rates have increased. This is again due to a mismatch between the original signal and the PR4 target response, which is why the PR4 shaping filter introduces correlation in the noise. PR4 (C2) still outperforms the other two algorithms, showing the value of exploiting the correlation across signal samples.

FIG. 10 shows the error rates obtained when using the EPR4 detectors. Due to a higher density, the media noise is higher than in the previous example with symbol separations of 4.4. This is why the graph in FIG. 10 has moved to the right by 2 dB in comparison to the graph in FIG. 8. While the required S(AWG)NR increased, the margin between the EPR4(Luc) and EPR4(C2) also increased from about 0.5 dB to about 1 dB, suggesting that the correlation-sensitive metric is more resilient to density increase. This is illustrated in FIG. 11 where the S(AWG)NR required for an error rate of 10^{-7} is plotted versus the linear density for the three EPR4 detectors. From FIG. 11 it can be seen that, for example, with an S(AWG)NR of 15 dB, the EPR4(Luc) detector operates at a linear density of about 2.2 symbols/PW50 and the EPR4(C2) detector operates at 2.4 symbols/PW50, thus achieving a gain of about 10% of linear density. Symbol separation of 2.9. This recording density corresponds to a symbol density of 3 symbols/PW50. Due to a very low number of symbols per a, this is the density where the detectors significantly lose performance due to the percolation of magnetic domains, also referred to as non-linear amplitude loss or partial signal erasure. FIGS. 12 and 13 show the performance of the PR4 and EPR4 families of detectors at this density. The detectors with the C2 metric outperform the other two metrics. The error rates are quite high in all cases. This is because at the symbol separations of 2.9, nonlinear effects, such as partial erasure due to percolation of domains, start to dominate. These effects can only be undone with a nonlinear pulse shaping filter, which have not been employed here.

The experimental evidence shows that the correlation sensitive sequence detector outperforms the correlation insensitive detectors. It has also been demonstrated that the performance margin between the correlation sensitive and the correlation insensitive detectors grows with the recording density. In other words, the performance of the correlation insensitive detector deteriorates faster than the performance of the correlation sensitive detector. Quantitatively, this margin depends on the amount of correlation in the noise passed through the system. Qualitatively, the higher the correlation between the noise samples, the greater will be the margin between the CS-SD and its correlation insensitive counterpart.

While the present invention has been described in conjunction with preferred embodiments thereof, many modifications and variations will be apparent to those of ordinary skill in the art. For example, the present invention may be used to detect a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel. The foregoing description and the following claims are intended to cover all such modifications and variations.

What is claimed is:

1. A method of determining branch metric values for branches of a trellis for a Viterbi-like detector, comprising: selecting a branch metric function for each of the branches at a certain time index; and applying each of said selected functions to a plurality of signal samples to determine the metric value corresponding to the branch for which the applied branch

US 6,201,839 B1

14

metric function was selected, wherein

corresponds to a different sampling time

2. The method of claim 1 further comprising

receiving said signal samples, said signal set

signal-dependent noise, correlated noise, or

dependent and correlated noise associated therewith

3. The method of claim 1 wherein said branch

functions for each of the branches are selected

signal-dependent branch metric functions.

4. A method of determining branch metric

branches of a trellis for a Viterbi-like detector,

selecting a branch metric function for

branches at a certain time index from a

dependent branch metric functions; and

applying each of said selected functions to

signal samples to determine the metric

corresponding to the branch for which the ap

metric function was selected, wherein

corresponds to a different sampling time

5. The method of claim 4 further comprising

receiving said signal samples, said signal set

signal-dependent noise, correlated noise, or

dependent and correlated noise associated therewith

6. A method of generating a signal-depen

weight for branches of a trellis for a Viterbi-

comprising:

selecting a plurality of signal samples, w

sample corresponds to a different sam

instant;

calculating a first value representing a bea

joint probability density function of a s

signal samples;

calculating a second value representing a

dent joint probability density function

samples;

calculating the branch weight from said

values; and

outputting the branch weight.

7. The method of claim 6 further comp

correcting the branch weight by an addit

8. The method of claim 6 further comp

correcting the branch weight by a multi

9. The method of claim 7 wherein

includes the step of selecting a third

prior branch probability for use as said

10. A method of generating a branch

of a trellis for a Viterbi-like detector, w

used in a system having Gaussian no

selecting a plurality of signal sa

sample corresponds to a dif

instant;

calculating a first value represen

quotient of a determinant of a

covariance matrix of said sig

minant of a trellis branch dep

in cov

of a subset of said signal samp

as

calculating a second value represen

a qu

signal samples less a pluralit

ized by a trellis branch depend

covariance of said subset of

signal samples;

calculating a third value represent

a quadratic of a

subset of said signal samples less

a plurality of channel

target values normalized by a

trellis branch dependent

covariance of said subset of

signal samples;

calculating the branch weight

from said first, second, and

third values; and

10. A method of generating a branch weight for branches of a trellis for a Viterbi-like detector, wherein the detector is used in a system having Gaussian noise, comprising:

selecting a plurality of signal samples, wherein each sample corresponds to a different sampling time instant;

calculating a first value representing a logarithm of a quotient of a determinant of a trellis branch dependent covariance matrix of said signal samples and a determinant of a trellis branch dependent covariance matrix of a subset of said signal samples;

calculating a second value representing a quadratic of said signal samples less a plurality of target values normalized by a trellis branch dependent covariance of said signal samples;

calculating a third value representing a quadratic of a subset of said signal samples less a plurality of channel target values normalized by a trellis branch dependent covariance of said subset of signal samples;

calculating the branch weight from said first, second, and third values; and

'839 Patent at Claim 10

Specification

• Covariance used in mathematical calculations

correspond to these samples. They are used to compute the vector \underline{N}_t , with which the empirical rank-one covariance matrix $\underline{N}_t \underline{N}_t^T$ is formed. In the absence of prior information, this rank-one matrix is an estimate for the covariance matrix for the detected symbols. In a recursive adaptive scheme, this rank-one data covariance estimate is used to update the current estimate of the covariance matrix $\hat{C}(\hat{a})$. A simple way to achieve this is provided by the recursive least-squares (RLS) algorithm. The RLS computes the next covariance matrix estimate $\hat{C}(\hat{a})$ as:

$$\hat{C}(\hat{a}) = \beta(t) \hat{C}(\hat{a}) + [1 - \beta(t)] \underline{N}_t \underline{N}_t^T$$

'839 Patent 9:45-56

estimation of the... Assume that the... observed. Based on the... after an appropriate delay... a decision is made that the most likely... for the sequence of symbols $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{L+1}$ is $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{L+1}$. Here L is the noise correlation length and $K=K_0+K_0+1$ is the ISI length. Let the current estimate for the $(L+1) \times (L+1)$ covariance matrix corresponding to the sequence of symbols $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{L+1}$ be $\hat{C}(\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{L+1})$. This symbol is abbreviated with the shorter notation, $\hat{C}(\hat{a})$. If the estimate is unbiased, the expected value of the estimate is:

$$\hat{C}(\hat{a}) = E[\underline{N}_t \underline{N}_t^T] \quad (21)$$

where \underline{N}_t is the vector of differences between the observed samples and their expected values, as defined in (12). Note that once the samples $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{L+1}$ are observed, and once it is decided that most likely they resulted from a series of written symbols $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_{L+1}$, the sequence of target (mean) values m_1, m_2, \dots, m_{L+1} is known that correspond to these samples. They are used to compute the vector \underline{N}_t , with which the empirical rank-one covariance matrix $\underline{N}_t \underline{N}_t^T$ is formed. In the absence of prior information, this rank-one matrix is an estimate for the covariance matrix for the detected symbols. In a recursive adaptive scheme, this rank-one data covariance estimate is used to update the current estimate of the covariance matrix $\hat{C}(\hat{a})$. A simple way to achieve this is provided by the recursive least-squares (RLS) algorithm. The RLS computes the next covariance matrix estimate $\hat{C}(\hat{a})$ as:

$$\hat{C}(\hat{a}) = \beta(t) \hat{C}(\hat{a}) + [1 - \beta(t)] \underline{N}_t \underline{N}_t^T \quad (22)$$

Here, $\beta(t)$, $0 < \beta(t) < 1$, is a forgetting factor. The dependence on t signifies that β is a function of time. Equation (22) can be viewed as a weighted averaging algorithm, where the data sample covariance $\underline{N}_t \underline{N}_t^T$ is weighted by the factor $[1 - \beta(t)]$, while the previous estimate is weighted by $\beta(t)$. The choice of $\beta(t)$ should reflect the nonstationarity degree of the noise. For example, if the nonstationarity is small, $\beta(t)$ should be close to 1, while it should drop as the nonstationarity level increases. The forgetting factor is typically taken time-dependent to account for the start-up conditions of the RLS algorithm in (22). As more data is processed, a steady-state

$\beta(t)$ is made to approach a close to zero, to reflect the (3), and to rely more on the increased and settles around

Estimates in (22) decays exponentially (22) can be started with covariance matrix $\hat{C}(\hat{a})$, with at the matrix be positive or an identity matrix.

Suppose that, prior to making the decision $\{\hat{a}_t, \hat{a}_{t+1}, \hat{a}_{t+2}\} = \{\Theta, +, -\}$, the estimate for the covariance matrix associated with this sequence of three symbols is

$$\hat{C}(\Theta, +, -) = \begin{bmatrix} 0.5 & -0.2 \\ -0.2 & 0.8 \end{bmatrix} \quad (24)$$

Let the forgetting factor be $B=0.95$. To update the covariance matrix the vector is first formed:

$$\underline{N} = (r_t - 1)(r_{t+1} - 0)^T = [-0.1 \ -0.2]^T \quad (25)$$

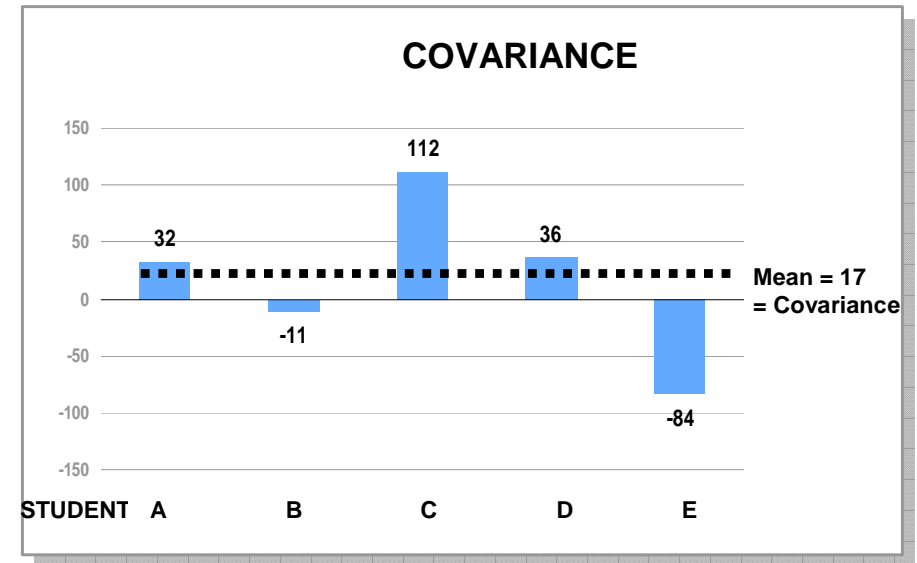
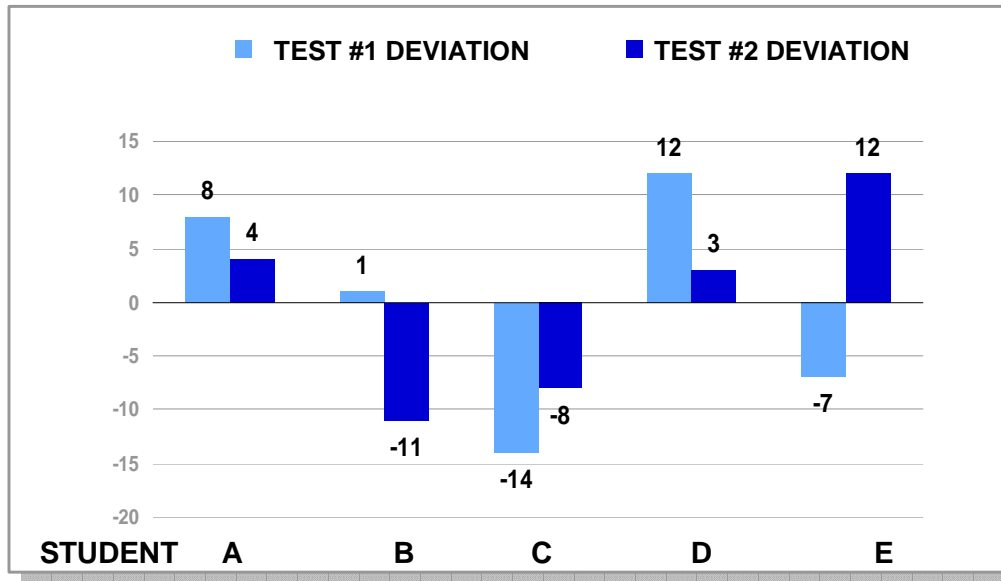
The rank-one sample covariance matrix $\underline{N} \underline{N}^T$ is used to find the covariance matrix update:

'839 Patent 10:53-67

Background: Statistical Meaning

(Marvell Tutorial Slide 64)

COVARIANCE the average of the pairwise products of test score deviations.



	TEST #1 SCORE	TEST #1 MINUS MEAN	TEST #1 DEVIATION SQUARED	TEST #2 SCORE	TEST #2 MINUS MEAN	TEST #1 x TEST #2 PRODUCT	TEST #1 DEVIATION x TEST #2 DEVIATION
STUDENT A	92	8	64	87	4	8004	32
STUDENT B	85	1	1	72	-11	6120	-11
STUDENT C	70	-14	196	75	-8	5250	112
STUDENT D	96	12	144	86	3	8256	36
STUDENT E	77	-7	49	95	12	7315	-84
	SUM 420		SUM 454	SUM 415		SUM 34945	SUM 85
	÷5 84		÷5 91	÷5 83		÷5 6989	÷5 17
	MEAN	DEVIATION	VARIANCE	MEAN	DEVIATION	CORRELATION	COVARIANCE

Extrinsic Evidence: Technical Treatises

- Marvell construction identical to ordinary meaning
 - ▶ “the expected value of the product of the deviations of two random variables from their respective means.”

The second joint *centralized* moment is called the *covariance*, and is denoted by K_{XY} :

$$K_{XY} \triangleq E\{(X - m_X)(Y - m_Y)\}. \quad (2.21)$$

Gardner, *Introduction to Random Processes with Applications to Signals and Systems*, at 32 (1986) (Marvell Exh. 24)

The *covariance* of two random variables X and Y , written $\text{Cov}(X, Y)$ is defined by

$$\text{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

where μ_x and μ_y are the means of X and Y , respectively.

Ross, *Introduction to Probability and Statistics for Engineers and Scientists*, at 119 (2000)

- ▶ See also Proakis Decl. at ¶¶ 34-35. (Marvell Exh. 17)
- ▶ Same definition in technical treatises and general dictionaries

co-vari-ance \,kō-'ver-ē-ən(t)s, -'var-; 'kō-,\ n (1931) 1: the expected value of the product of the deviations of two random variables from their respective means 2: the arithmetic mean of the products of the deviations of corresponding values of two quantitative variables from their respective means

Merriam-Webster's Collegiate Dictionary (10th ed. 1995) (Marvell Exh. 15)